60th NORTHEAST REGIONAL STOCK ASSESSMENT REVIEW COMMITTEE (SARC-60)

Review report for the benchmark stock assessment for scup and bluefish, SAW/SARC60

Sven Kupschus

Prepared for

Center for Independent Experts

The Centre for Fisheries and Aquaculture Science
Lowestoft Laboratory
Pakefield Road
Lowestoft
Suffolk NR33 0HT
England, United Kingdom

Phone: +44 1502 524454
Email: sven.kupschus@cefas.co.uk
www.cefas.co.uk

Contents

1. Background

The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Representative (COR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are independently selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

2. Review activities

The Review Committee convened at the Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from June 2nd -5th 2015. The Committee comprised a chair and three panel members. Plenary sessions were open to the public.

A formal presentation of the Stock Assessment Workshop (SAW) results was given by the lead assessors from each of the working groups (scup: Mark Terceiro, bluefish: Tony Wood) and specific issues were discussed. The assessors returned, when required, for further discussion and clarification of how the SAW Terms of Reference were addressed, including carrying out some additional model runs for clarification or sensitivity tests.

The panel members were then required to prepare an independent report indicating for each Term of Reference of the relevant SAW: i) whether the work that was presented is acceptable based on scientific criteria (e.g., consider whether the data were adequate and used properly, the analyses and models were carried out correctly, and whether the conclusions are correct/reasonable); and ii) whether the work provides a scientifically credible basis for developing fishery management advice.

The SARC chair and panel members prepared a first draft of the consensus report during the meeting. The panel members prepared their independent reports following the meeting. There was general consensus amongst the panel regarding the completion of ToRs by the WG and the status estimation and projections of scup or bluefish. Therefore my independent review presented below to a large extent reflects the consensus report developed at the meeting, with additional comments. Some of the original Consensus Report text has been summarised, or expanded where appropriate, but without changing the Committee's agreed views.

3. Acknowledgements

I would like to thank all the technical working group members contributing to the meeting for their informative presentations of the SAW results and for providing helpful response to the

SARC's questions. My appreciation also to staff at the Woods Hole Laboratory and particularly to Jim Weinberg, Sheena Steiner and Chris Legault for their hospitality and help throughout the meeting. Lastly, I extend thanks also to the other members of SARC for productive discussions on the assessments and Cynthia Jones as chair for keeping the necessary focus of the meeting.

4. Executive summary by stock

Scup

I agree with the WG that the scup stock is not overfished and that overfishing is not occurring. The accepted assessment seems to reflect the general temporal trends in F and SSB, but I found it hard to accept the extreme rate of decline in F to recently very low levels. The cause of the poor estimation of the decline in F and subsequent expansion of SSB seems to be the inclusion of the historic catch data in conjunction with model assumptions.

Stock status estimation and stock projection based on the model appear to be robust to the weaknesses in the assessment because management metrics and biomass reference points appear to be equally affected by alternate views. Absolute values of current F and SSB are much more uncertain than the ratios of SSB / $B_{MSYproxy}$ and F / $F_{MSYproxy}$.

Bluefish

I agree with the WG that the bluefish stock is not overfished and that overfishing is unlikely to be occurring. The accepted assessment presented a thorough review of the data and modelling procedure for what is a difficult stock to model given the migratory behavior and spatio-temporal variability in gears used to exploit them. The main weakness remaining in the assessment is the heavy reliance on the MRFSS/MRIP CPUE index. The index is not independent of the catch data, sharing substantial information with 80% of catches coming from the recreational sector covered. In addition MRFSS and MRIP data are generally thought to be associated with high variance estimates. No indication was found that either bias or variance of the index negatively affected the assessment and no other data on older specimens are available to replace the index.

Central tendencies of stock status estimation and stock projections of the model appeared to be robust but variability, with only a 15% inflation factor over the CV coming out of the assessment, suggests that the likelihood of the stock being overfished is possibly a little greater than the 0% suggested by the current projections.

5. Assessment of scup

Summary

I agree with the WG that the stock is not overfished and that overfishing is not occurring. The assessment as presented seems to reflect the general trends in F and SSB, but I found it hard to accept the extreme rate of decline in F to recently very low levels. The cause of the poor estimation of the decline in F and subsequent expansion of SSB seems to be the inclusion of the historic catch data in conjunction with model assumptions.

Stock status estimation and stock projection based on the model appear to be robust to the weaknesses in the assessment, because management metrics and biomass reference points appear to be equally affected by alternate views. Absolute values of current F and SSB are much more uncertain than the ratios of SSB / $B_{MSYproxy}$ and F / $F_{MSYproxy}$.

Results by TOR

1. **Estimate catch from all sources including landings and discards. Include recreational discards, as appropriate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.**

This ToR was addressed successfully by the WG: Seen in isolation the derivation of the catch data is appropriate, but a more detailed analysis is only possible within the assessment framework as the interpretation of the data intrinsically depends on how the assessment model deals with the information in the context of the assumptions.

The presentation provided by the WG described the derivation of the catch (landed and discarded portions) based on previous methodology developed for the stock. This meant relatively little information was available on the detailed derivation of data and availability of current data sources for this assessment. I do not question the data sources used or the data derivation methodology, and trust that the WG, being much more closely involved with the data, are best placed to make the appropriate decisions. I do, however, think that this information could be presented to reviewers in a more readily digestible format without having to search through historic documents to find out the reasoning behind certain decisions.

One major change in estimating commercial discards was the change from a species specific discard ratio to an all catch based discard ratio stratified by gear and quarter (MESH240). The newly recommended discard estimates seem to combine the historic perception (discard high pre-1994, low 1994-2005) common to all MESH indicators with the more recent view of higher total discards in the recent period (2006 – present) demonstrated in the GMDL method (with higher sampling rates). Although I felt that this methodology could be supported based on knowledge of the fisheries, the sampling effort and the statistical properties of ratio estimators, I was unable to follow the working group's reasoning for this choice in detail. The working document states "comparison with the current GMDL method estimates indicates that the SBRM MESH240 estimator and stratification provides the best overall combination of feasible estimates of the scup discards and landings and good precision", but does not provide the details as to the definition of "feasibility" and "good precision". The way the 1998 data is dealt with in light of one sample with very high leverage is a little questionable. Discard data seem to be highly influential in estimating recruitment in the model since a large proportion of discards is at this age. Averaging adjacent year discard estimates in the way described in the working paper seems to me to run the risk of 'smearing' recruitment signals in this period and potentially could have impacts on the estimates of catchability for the entire time block. One potential way to handle this in a SCA model would be to reduce the effective sample size in the final assessment to reflect the greater uncertainty. The same criticism applies to the estimation of historic discards pre-1989 where average discard ratios 1989-2001 were applied.

Ageing, natural mortality and maturity:

Given the reliance on age information, both in terms of catches and aged index information, the accuracy of aging data is critical to the assessment. The estimation of the accuracy of the available age data seems to be heavily based on the ability of different readers to replicate age readings, and much less on whether the ages thus derived are true representation of the ages sampled. Because scales rather than otoliths are used, it is possible that older ages are systematically underestimated as has been found for other species. The likelihood of such biased aging errors is difficult to determine in this assessment due to the generally large variability in the abundance-at-age data. Surveys in

particular are characterized by year and age effects (based on the residual pattern of the assessment) rather than by cohort effects making it difficult to discern systematic ageing bias at older ages. I conducted a brief examination of the assessment residuals for signs of spreading of cohorts or jumping between cohorts of high or low abundance. Invariably it seemed that the variability in abundance-at-age information was considerably larger than any potential aging bias, but one cannot conclude from this that no aging bias exists in the data. Catch weight-at-age also indicated that variability was characterised by annual effects with all ages increasing or decreasing in synchrony. As growth effects are cumulative they should be more linked to cohort than age effects. The absence of such effects in the data suggest that sample variability in general is greater than the size of these effects and that any possible effect of aging or density dependent growth are likely masked by this variability. Given the variability in age information it seems unlikely that this data could allow for the estimation of natural mortality within the assessment, therefore the M=0.2 assumption based on life history traits seems appropriate. Maturity information has been revised as part of this assessment. However, given the concerns over age information it is not clear to me whether this represents a better estimate of maturity or not. However, the final assessment does not imply a stock recruitment pattern so that the state of the stock assessment is largely robust to the assumption about maturity with both the current SSB estimate and the SSB reference point scaled by the same maturity information.

The overall impression of the data derivation to me seems like it is based on sound decision making and the comments regarding the decision making are more about how the data are presented rather than the data themselves. I therefore determine the term of reference to have been met by the WG but feel that the characterization of uncertainty could be improved through better presentation of the basis of the decision making process.

2. **Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). Characterize the uncertainty and any bias in these sources of data.**

This ToR was addressed successfully by the WG: A large number of survey indices were available to the WG. Most were relatively local and short term in nature and represented divergent views of the dynamics of the stock, likely due to spatio-temporal variability in the distribution of the population at finer spatial scales than the federally based longer-term time series (NEFSC spring and fall surveys). Only some indices provided information on the age composition, and then predominantly on younger ages. It is difficult to summarise this information in a coherent way and the WG tried to do so by providing different means of combining the information into a single index both as a tool to describe the survey information and for inclusion in potential future assessments.

I appreciated the presentation of this information on both the GLM and hierarchical approaches. The aim of these methods is to be objective in the extraction of information. Conn (2010) introduces and tests the later methodology. However, in the case of the scup data, where there is relatively little spatial overlap in the survey indices, it seems unlikely that the methodology is capable of ascertaining the relative proportions of the population found in each area while simultaneously estimating catchabilities.

Both methods are able to be applied to age based data, though in this case this was only done for the GLM approach. It is unclear as to how the model could adequately cope with potentially different selectivities/catchabilities for different surveys when dealing with age-aggregated data (CPUE). As an independent method to evaluate the utility of certain indices in order to reduce the complexity of the assessment model, I can see that this approach has merit (but only when conducted by age). It does not, however, remove the subjectivity of the choice of indices to include in the final model as there is no demonstrably useful cut-off criteria that can be invoked to make that choice on an objective basis. As a means of reducing the multiple indices to a single index for inclusion on the assessment model, it is difficult to see for either method (GLM or hierarchical) how it could work. If the estimation of the selectivity and catchability parameters is carried out external to the assessment model, it is difficult to see how the uncertainty information on these parameters could be effectively transferred to an

assessment model dealing with a wider set of information sources. In other words, the proposed combined index is conditional on the specific estimation of selectivity and catchability. If the stock assessment on the basis of other data sources were to estimate grossly different estimates of selectivity/catchability of the index information, then the aggregated index no longer provides an unbiased estimate of the population. If parameter estimates were similar then almost certainly the assessment model would have drawn the same conclusions regarding the individual surveys as the external model, so nothing would have been gained by the prior external estimation.

Despite the limited value of the above methodologies in the formal assessment process, the information presented did allow for a good summary of the survey information and a number of important conclusions about the relevance of individual indices could be made. Generally speaking, spring surveys were more responsive to recruitment, while fall surveys tended to be more informative on the older age classes. Large scale surveys did prove in most cases to be more consistent estimators of population abundance than smaller scale surveys. Final choice of index inclusion was made within the assessment model itself largely based on the evaluation of the root-mean-square estimators. The final choice of indices seems to reflect the information content of the individual surveys.

3. **Describe the thermal habitat and its influence on the distribution and abundance of scup, and attempt to integrate the results into the stock assessment.**

This ToR was addressed successfully by the WG: The presented information indicated that the habitat model could be successfully applied to determine the effects of environmental conditions on the availability of scup to survey catches. Unfortunately, the variability in availability was small and mostly random throughout the time period of the surveys. It was hoped that the approach would explain spatial differences in the distribution of scup in some survey years that had led to substantial year effects in the NEFSC surveys. A few large catches of scup were encountered in large strata not usually occupied by the species, but sampled regularly for other species. Multiplying these catches by a large areas compared to the size of strata that scup are usually found meant that these relatively small number of stations carried much greater leverage in the index calculation in those years than in others.

It is not clear from this analysis whether the distribution of scup is spatially affected by factors other than temperature, or if these spikes are stochastic and better methods of dealing with a more contagious error distribution within the model need to be sought in order to reduce the leverage of such stations in the final model.

The thermal niche dependent survey availability approach has been employed in the assessment process for butterfish. I am unable to comment on the justification in that assessment. The attempt to apply this to the scup assessment (using surface and bottom salinity and temperature) when the evidence for a high degree of temperature or salinity dependence in scup distribution seems to be lacking, gives the impression of a cookie-cutter approach. The risk of such approaches generally is the same for the correlative application of a large number of independent variables, some (5% at alpha=0.05) of which may be correlated merely by chance. In this case it did not solve the intended data issues for the assessment and given the difficulty in obtaining the most recent hind casts for environmental information the correction was not used. A more process focused approach, i.e., determining the variables important in the distribution of a species and then using these in similar model, is likely to provide more robust assessment estimates.

4. **Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.**

This ToR was addressed successfully by the WG: In general the assessment seems to reflect the trends in the stock dynamics on a relative basis. I am of the opinion that the sizable decline in catches is overly interpreted as a very steep decline in F rather than a combination of reduced recruitment and

declining F. The consequence of this is that the model is forced to estimate F to be very low in the period 2005 onwards with a commensurate overly optimistic trend in SSB over the same period. However, alternate model configurations are unlikely to materially alter the perception of a generally increasing biomass under more or less constant F over the period. Hence, stock status estimations from alternate assessments are unlikely to indicate that overfishing is occurring in the stock.

Whether the stock is overfished would be difficult to determine on an absolute basis. Here it is based on an SPR ratio mainly because current assessment provides little evidence for a stock-recruit relationship. The SPR approach is more robust to inaccuracies regarding the absolute values of SSB than values based on spawner-recruit assumptions. Although in my opinion not the best possible model configuration, the base model presented by the WG is suitable for the provision of advice in the short-term, given the current development of stock trends. A persistent rise in F, or decrease in recruitment sufficient to reverse the generally increasing trend in SSB, would substantially reduce the confidence in the assessment of stock status based on the current assessment. Such future stock development would mean the stock status estimates based on the current model would be overly optimistic.

During the SARC meeting a number of alternate model configurations were presented either to assess sensitivity to certain assumptions or as requests to the WG for additional information. On the basis of these additional analyses, I developed my conclusions regarding the optimistic interpretation of the data by the base assessment. Although I believe that better model configurations can produce more realistic assessments of recent stock trends, it does not undermine the suitability of the current BASE_18 model to produce management advice given the earlier described caveats. What follows is a brief summary of the reasons for those conclusions.

1) The BASE_18 model is surprisingly insensitive to changes in the model setting over the central period of the assessment time series. This might be interpreted as a positive characteristic, but introduction of unreasonable assumptions, for example high and low M estimates, lead to very similar F estimates compared to the base model. More often one would expect changes to a single M value to have scaling effects on F and recruitment across the entire time series. In addition, the MCMC examination suggests that historic biomass levels are far more accurately defined when the available data were characterized as poor and age information was unavailable, while current estimates are much more uncertain despite much better data. The systematic underestimation of historic uncertainty is more commonly associated with VPA type models, something an SCA model such as ASAP is specifically designed to redress. Lastly, changes to the model configuration suggest some dramatic and unrealistic alternatives to the BASE_18 view for this early part of the time series.
2) The retrospective pattern in F shows dramatic changes in the historic estimates of F, particularly in 1968, while recent estimates are highly consistent in absolute terms. In contrast, recent recruitment is almost always underestimated while historic recruitment (1963-1983) is estimated to have been consistently higher for later assessments. The intervening period 1984-2003 is very stable in the retrospective analysis both in terms of recruitment and F. This suggests to me that the model is fixed in this central period, where in all retro-models F first increases steadily and then drops sharply. Any movement in the absolute values is restricted in this period so that the model, rather than scaling up or down as expected, flexes around this period and exaggerates the changes in the early and late period where it is more free to do so.
3) The sensitivity to mortality demonstrated a similar effect in terms of the response in recruitment. The central period was relatively unaffected while absolute estimates of R showed a greater range in the early and late periods. In contrast, apart from the early spike in F, said to be caused by the initial parameter estimates, most of the variation in F under different M assumptions was found in the central period.
4) Altering the shape of the commercial catch selectivity curve (dome shaped to flat topped) produced surprisingly large effects in the absolute estimates of final year F and SSB (around 40%). Despite this the overall trends in stock dynamics were persistent, merely the rate of

decrease in F 1994-2002 decreased and the stock expansion since this time was slower given the higher Fs. From a statistical perspective, there was little to choose between the two selectivity models. The residual patterns for the catch were marginally improved at the older ages and the penalty function increased in the model as expected given that it fit 13 less parameters. It was interesting to see though that a substantial part of the increase in the penalty came from the deviates from the initial parameter estimates, in other words, the starting values for the initial population numbers. The starting values for these were more or less randomly chosen and it was anticipated the deviates from the initial values would free up the model. I am not convinced that this is working effectively for scup.

5) The BASE_18 model suggested a near 20 fold decline in F over the period 1994 to 2002. Such a dramatic decline in F was not discernable in the commercial effort data. Although the effort data are considered uncertain, neither this, nor the potential effect of the recently imposed GRAs, thought to mainly affect very young scup, can explain why such a dramatic decline is not apparent in the effort data. Logbook (WP5) and VTR (WP6) data respectively report declines in days fished of 3 or 4 fold over that period for the otter trawl gear responsible for most of the commercial catches. Even taking into account the uncertainty of the estimates and the ability for fishermen to change behavior, I find it hard to reconcile the dramatic decline in F with the much less variable effort data.
6) A sensitivity to the length of the time series used in the BASE_18 assessment indicates that a starting date of 1989 commensurate with the initial age collections for the catch data favors a much reduced decline in F with substantially greater estimates of current F and 40% less terminal SSB. Although there is value in extending the time series back as far as possible to give a better historical perspective, I cannot see the justification for this when there is so little evidence that this historic period is driven by signals in the data. Some form of recruitment signal is necessary to inform on changes in F. Although one might argue that the discarded biomass might contain significant information on recruitment, it has in fact been derived from the average proportion of catches discarded in a later period and provides no cohort signal. Consequently, it is not possible to estimate recruitment deviates for this period (lack of model convergence or unrealistic results). The model is forced to accept average recruitment in the post 1989 period and cannot vary the scale of recruitment. Given the autocorrelation seen in more recent recruitment and the auto correlated nature of historic landings (back to 1940) it seems highly unlikely that such an assumption is appropriate for this stock.

My conclusions from these limited analyses is that the rate of decrease in F over the period 1994-2002 is exaggerated, and that this is at least in part caused by the assumptions about the initial population in conjunction with the lack of information on ages in the early part of the time series. Shorter time series suggest that current F is higher and current SSB is lower than the BASE_18 model. The initial population assumptions also force the model into accepting a dome shaped selectivity for which there appears to be little evidence in the data itself, but which produce an (in my opinion) overly optimistic view of SSB trends.

The historic perspective would be informative for management, but in my opinion it does not reflect the stock dynamics in that period as indicated by the sharp spike in F (1968) following the initiation of the population. The argument to view this as a 'burn in' period is counter intuitive a) because the idea is of a historic perspective which is then not used, and b) because the assumptions have considerable effects on stock trends in the later period as shown in the retrospective and M sensitivity runs. Finally, the choice of 1964 as a start date seems somewhat arbitrary when catch data exist back to 1940. Catches at that time appear to have been considerably higher.

5. **State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment**

on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

This ToR was addressed successfully by the WG: The existing status definitions for 'overfished' and 'overfishing' for scup are based on recommendations from the 2008 Data-Poor Stocks Working Group (DPSWG) Peer Review Panel (NEFSC 2009). For scup a biomass reference point of 40% SPR, is recommended with a default threshold fishing mortality reference point that maintains SSB at 40% SPR. The WG stated the existing definitions and values for $F_{MSYproxy}$ = F40%, $B_{MSYproxy}$ = SSB40%, $B_{THRESHOLD}$ = ½ SSB40%, and MSY.

Based on the BASE_18 ASAP assessment the WG quantified the point estimates of these reference points as follows:

$F_{MSYproxy}$ = F40% = 0.220 > 0.127 = F_{2014}
$B_{MSYproxy}$ = SSB40% = 87,302 mt < SSB_{2014}= 182,915 mt
$B_{THRESHOLD}$ = ½ SSB40% = 43,651 mt
MSY = MSY40% = 11,752 mt
(comprised of landings of 9,445 mt and discards of 2,307 mt)

The examination of the stock-recruit relationship for scup indicated that the even at the low biomass predicted by the assessment there was no clear evidence that the reproductive capacity of the stock had been compromised. Rather there appeared to be some temporal auto correlation in the recruitment, and that current recruitment levels appeared to higher than under the previous regime ending in the late 1990s. Given this the WG moved to the estimation of biomass reference point proxies based on per-recruit calculations as advised by the DPSWG in 2008. I consider this approach appropriate.

Uncertainty estimates of the reference points were not provided. While it is possible to derive uncertainty estimates for the reference points based on the assessment based on the MCMC evaluation, this was not examined by the WG, or the results were not presented. However, the uncertainty in the reference points is intricately linked to the uncertainty in the assessment. Although the absolute values of the reference points are uncertain, the real quantities of interest are the ratios of $F/F_{MSYproxy}$ and $SSB/B_{MSYproxy}$ which are likely to be highly conservative. Consequently, the panel did not consider it necessary to specifically examine the uncertainty in the absolute level of the reference points and accepted the approach taken and values provided by the WG.

6. **Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review.**
   a. **When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.**
   b. **Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).**

This ToR was addressed successfully by the WG:

a. The existing model updated with new data indicated that the scup stock was not overfished and overfishing was not occurring in 2014 relative to the existing (old) biological reference points established in the 2008 Northeast Data Poor Stocks Working Group (DPSWG; NEFSC 2009) assessment. The F was estimated to be 0.049 in 2014, below the fishing mortality threshold = $F_{MSYproxy}$ = F40% = 0.177. Spawning Stock Biomass (SSB) was estimated to be 219,066 mt in 2014, above the biomass target reference point = $SSB_{MSYproxy}$ = SSB40% = 92,044 mt.
b. The panel accepted the new ASAP model recommended by the SAW as a basis for providing status determination with some caveats (see ToR4 in the summary report elaborated on in this report under the same ToR). Having accepted the model (ToR 4) and the biological reference

points (ToR 5), one concludes that the stock is currently not overfished and that overfishing is not occurring. See ToR5 for estimates from the accepted assessment.

7. **Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) (see Appendix to SAW TORs for definitions).**

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

a. Projections under two assumptions for the interim year catch ($catch_{2015} = ACL_{2015}$, $catch_{2015} = 0.75 * ACL_{2015}$) were provided by the WG using the 2010-2014 patterns of discarding to split catches into landings and discards. The ToR calls for the provision of annual probabilities of exceeding threshold reference points (below for SSB and above for F); these are not specifically shown by the WG report but it is clear from the projected estimates of SSB and from the text that the probabilities in each of the years 2016-2018 of exceeding the biomass threshold is zero. Similarly, the probability of exceeding the F threshold is very low, if not zero. The ToR calls for a sensitivity analysis in which a range of assumptions about the most important uncertainties in the assessment are considered. The WG has stated that biological inputs to the scup assessment are well-founded, and has attended to process error and retrospective errors in terminal year estimates of abundance by inflating (doubling) the OFL CV used in the projection. This was deemed sufficient to account for the uncertainty seen in final year abundance/SSB estimates across a wide range of sensitivity tests carried out by the WG and the uncertainty estimates covered a number of sensitivities explored, such as M and different selectivity patterns.

b. The WG provided catch options under the two assumptions of catch in 2015, but preferred the option that assumes the ratio between landings and ACL remains constant. This seems reasonable, although a better alternative might be to assume that the $F_{2014} = F_{2015}$ as this would be more responsive to changes in the population. Other uncertainties were covered under the inflated CV implemented in the OFL calculation.

c. The SAW has attended to this ToR with respect to landings compared to MSY and historical values. This addresses the productivity aspect of the ToR, but does not consider the slightly wider scope of vulnerability. The species is not thought to be under environmental stress, but historically landings have been considerably higher periodically. Whether such changes are indirectly human induced or a response to random environmental fluctuations is not clear. There is no substantiated evidence that the fishery creates susceptibility concerns, but for completeness comments on the susceptibility of scup should be included in future assessments.

8. **Review, evaluate and report on the status of the SARC, SSC, and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.**

This ToR was partially addressed by the WG: The progress against research recommendations was presented by the WG. The recommendations are set by various bodies, usually in response to previous assessments and their review. Frequently, they recommend a particular analyses or approach to be pursued, without specifically stating as to how this might be likely to improve the assessment in terms of its weaknesses. Some recommendations are based on reviews of other stocks, without a consideration of whether a similar approach may be useful for a specific species (for example, the

environmental correction of survey indices based on the available habitat niche). Others appear to relate to the expertise / beliefs of individual reviewers, who may well have a clear idea of how to use the recommended information, but it is not apparent how this fits in with the assessment process (for example conduct tagging).

In the absence of a better description it was difficult for the WG to do more than state progress. In order to provide more focused research in the future, the panel attempted to provide greater detail through their own interpretation in relation to the benefits to the assessment and the likelihood of success of each recommendation. It is hoped that this will make the process of prioritizing research easier and the improvement to assessments faster and better informed.

6. Assessment of bluefish

Summary

I agree with the WG that the stock is not overfished and that overfishing is unlikely to be occurring. The accepted assessment presented a thorough review of the data and modelling procedure for what is a difficult stock to model given the migratory behavior and spatio-temporal variability in gears used to exploit bluefish. The main weakness remaining in the assessment is the heavy reliance on the MRFSS/MRIP CPUE index. The index is not independent of the catch data, sharing substantial information with 80% of catches coming from the recreational sector. In addition, MRFSS and MRIP data are generally thought to be associated with high variance estimates. However, no indication was found that either bias or variance of the index negatively affected the assessment.

Central tendencies of stock status estimation and stock projections of the model appeared to be robust. When viewed independently, sensitivities of the assessment are relatively small, which led the WG to suggest that a 15% inflation factor of the CV on the estimate of the OFL is sufficient to represent the uncertainty within the projections. The additive sensitivity, although still small, is almost certain to be larger than maximal independent uncertainty. Consequently it is likely that the 0% probability of the stock being overfished given the recommended OFL is greater than the 0% suggested by the current projections.

Results by TOR

1. **Estimate catch from all sources including landings and discards. Evaluate and if necessary update the discard mortality estimate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.**

This ToR was addressed successfully by the WG:

Commercial Fishery Data:
Commercial landings have ranged from 20-37% of total landings with an overall declining trend. The commercial quota has not been taken since its introduction in 2000. Previously commercial landings have been estimated from the CFDBS administrated by the NEFSC with additional information supplied by individual states. For this assessment, the Atlantic Coastal Cooperative Statistics Program (ACCSP) database maintained by the Atlantic States Fishery Commission acted as a single source of information. However, a discrepancy of 1.5% was noted between the ACCSP Virginia commercial reported data and Virginia's Fishery Mandatory Reporting Program Trip (FSMRPT) landings database, and the Potomac River Fishery Commission

(PRFC). To deal with this discrepancy, the WG chose to use the greater of the landings recorded by the ACCSP and the VA historical landings database. The nature of the discrepancy, failure to load some records during migrations, suggests that the treatment is appropriate. The overall effect of potential alternate catch series was tested by sensitivity analyses and was minimal. Consistent quality information of discard rates over the time-series are available only for the trawl fishery, which now represents a small portion of the catches. The decision not to include commercial discard information in the assessment is justified by the low proportion of commercial landings and an overall small discard rate.

While bluefish were historically landed in the southeast and mid-Atlantic regions, most commercial harvest is now located in the mid-Atlantic region. Gillnets present the dominant portion of the catch

Recreational Fishery Data:
Recreational landings + dead discards (@ 15% hooking mortality) have ranged from 66-84% with an increasing trend. Data on recreational landings, releases, and biological metrics were obtained for 1981 through 2014 from NMFS MRFSS/MRIP. These surveys have two components: 1) an offsite survey to measure effort, and 2) an on-site access intercept survey to obtain CPUE and biological metrics. Total catch is then estimated by multiplying CPUE by effort. In 2005, results of a review of MRFSS by the National Research Council found bias in the access intercept portion of the survey. This bias was quantified through the MRIP beginning in 2004 and corrections applied to the historical MRFSS landing data. Sampling is stratified by state, mode of fishing, and wave (bi-monthly period). Catch is observed by the survey agent, but releases (discards) are self-reported by the angler.

The catch from the charter and for-hire recreational fishery has been surveyed through a separate program, the for-hire survey (FHS), since 2005. This provides a more precise estimate of this sector of the recreational fishery.

2. **Present and evaluate data and trends on life history information including, age, growth, natural mortality, food habits, and maturity.**

This ToR was addressed successfully by the WG: The comments below are a mixture of additional interpretation and more detailed context of the consensus report and reflect my own thoughts on the topics discussed.

Aging of bluefish for assessment purposes has become highly standardised and is now invariably based on otoliths because significant aging bias (under aging) has been demonstrated for scale samples. Some of the historic ages used in the assessment are still based on scales, and these data should generally be treated with a higher degree of uncertainty than more recent information. The WG conducted a full review of the available age information to characterize both the catch and the index information. Compared to previous assessments wider sources of information were used to derive semi-annual ALKs. Some historical scale information from North Carolina was corrected for the assumed January 1st birthday by adding a year to the recorded ages. This assumes that no part of the population has deposited an annulus prior to January first, and it is not entirely clear if this applicable for bluefish in those waters. However, the corrected information matched the length at age information from other areas much more consistently so it is likely that the correction applies to the majority of the aged individuals.

In a fast growing and migratory population there are risks in the correction and use of sources of age information with spatio-temporally varying sampling effort. In this case the benefits to the assessment in terms of greatly reducing the need to pool data across years, and thus reduce cohort signal in my opinion outweigh the risks of introducing some biases in the age composition due to spatially / temporally distinct growth rates.

Both commercial and recreational catches-at-age and MRFSS/MRIP CPUE data (using the same ALK but different length distributions) indicated cohort tracking, particularly in the historic data. This structure in the data is only possible if aging and relative abundance information are consistent. One concludes that despite the *a priori* concerns over aging from scales, use of aggregated ALKs and the precision of recreational catch statistics are not a major concern in this assessment. Aging bias from scales may be less of a problem at the ages used in this assessment as it is known to become more significant at higher ages and overlap in length between ages appears to be low.

Natural mortality for bluefish in the assessment is assumed to be 0.2 per year over all ages. Little evidence is available in literature specific to bluefish to suggest more appropriate values. Recent trends to determine M based on life history characteristics, including scaling mortality by age (or size) were also investigated, but produced estimates in broad agreement with the information used in the assessment. The only evidence that M may in fact be higher came from the understanding that a sizeable proportion of the stock may be taken by mako sharks each year. These estimates were based on a relatively small sample of mako stomachs, however, so are not sufficiently certain to warrant implementation of a higher M at this time.

Maturity information was reviewed in detail, taking account of spatial trends, sex and aging bias. The methodology employed in the end yielded very similar results to the most comprehensive, but localized study published in the literature, so it seems that the differences in sampling tend to average out over the population.

Data on food habits was presented for species consumed by bluefish. The species is very much a generalist having been shown to feed on whatever is available locally. Temporal variation in mortality and apparent changes in the size-at-age in bluefish are more likely to be linked to (size selective) predation rates and environmental factors than to prey availability post recruitment. Therefore, food habit studies for bluefish may be less of a priority that for more specialized species that are more likely to suffer from food limitation under specific ecosystem conditions.

3. **Present the survey data available for use in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.), evaluate the utility of the age-length key for use in stock assessment, and explore standardization of fishery-independent indices. Investigate the utility of recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data, including exploring environmentally driven changes in availability and related changes in size structure. Explore the spatial distribution of the stock over time, and whether there are consistent distributional shifts.**

This ToR was addressed successfully by the WG: The WG reviewed the available information describing standardised changes in the abundance of bluefish. Most of the surveys are localized and coastal and therefore contain information only on the youngest ages with little indication of the relative proportion of the population they represent. The highly migratory behavior of the species, with an incomplete picture of the proportions of the population that migrate at any one time, complicates the determination of a coherent picture within the assessment.

Attempts to produce a more coherent picture of recruitment information from these YOY-surveys external to the model using the hierarchical approach developed by Conn (2010) largely removed the contrast in recruitment pattern observed in individual surveys. In the absence of stronger, or more universally detectable variations in cohort strength it is not possible to evaluate the appropriateness of the approach for bluefish in more detail (See also comments on the methodology in the scup section of this report). The compound index for the seine surveys was used in the final assessment.

An attempt was made to apply the 'available habitat' survey correction method used in the butterfish assessment. Bluefish annual migrations are generally understood to be linked to temperature so that such an approach would be a useful measure to improve index precision. The correction was only applied to the NEFSC surveys where it explained little of the variation in the raw index. Applying a similar correction to spatially restricted state surveys might prove more effective as the proportion of the population outside the sampled areas is considerably larger. Given the small effect of the correction for the NEFSC surveys and the lack of recent hind cast environmental information, the methodology was not applied to the proposed assessment. An assessment sensitivity run was conducted suggesting minimal effect on management metrics.

Although not a survey, the MRFSS/MRIP information was used to develop an index that proved to be the only really influential index in the final assessment, largely because it provides substantial information on older ages. As described in the consensus report, the index was distinct from the recreational catch information in a number of ways, but relied in a substantial part on the same information and hence is not entirely independent of catches or discards from that fishery. The heavy reliance of the assessment on these data to describe changes in the population is undesirable, and priority should be given to developing fisheries independent source of information on older ages that can replace the MRFSS/MRIP information, or at least be used to quantify biases in those data. The use of the index in the assessment is a weakness of the approach taken, but given the lack of other usable information in conjunction with the extensive efforts to remove biases from the data, I consider the option the best available approach.

4. **Estimate relative fishing mortality, annual fishing mortality, recruitment, total abundance, and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Explore inclusion of multiple fleets in the model. Include both internal and historical retrospective analyses to allow a comparison with previous assessment results and previous projections. Explore alternative modeling approaches if feasible.**

This ToR was addressed successfully by the WG: The WG followed a thorough and strongly structured process in the development of the assessment model, heavily based on comments regarding the weakness of previous assessments while ensuring that remedial measures could be scientifically justified based on the available information. Consequently, I found the development of the new model easy to understand and decisions of options chosen justified. No doubt there are remaining weaknesses (mainly based on the data rather than the implementation) and some decisions where other assessors or reviewers (including myself) might have chosen different options. However, it seems the divergence of opinion is small because of the comprehensive foundation provided in the presentation, and moreover, it seems that they are largely inconsequential to the results of the assessment in terms of the trends in stock dynamics. For example, the minor misspecification in the model, though corrected in the final model (see consensus report for details), had little effect on the conclusions regarding stock status.

The model building process was successful in achieving the aim of reducing the constraints implemented in previous assessment approaches, which tended to employ subjective weighting of various data sources and therefore risked biasing assessment results. The success in developing a more balanced model is reflected in the balanced contribution of data sources to the penalty function, though some systematic patterns remain in the residuals suggesting that not all process errors were resolved.

There was some concern of the WG over the remaining retrospective pattern in the assessment and whether the criterion should be applied as a diagnostic of the assessment process (see comments in the consensus report). Given the divergent information in some of the data sources used, the temporal imbalance (short, recent time series that lose weight in the assessment as consecutive years are removed) and the low likelihood that the seperability assumption can be strictly applied to developing recreational fisheries, it seems to me to be improbable that the retrospective patterns could be resolved in anything but a highly complex model. The risk of the complex model

alternative would be over fitting, i.e., estimating processes for which we have no way of discerning whether the underlying considerations are justified. Taking this and the size of the retrospective bias in the B044 model compared to many other stocks into account, I feel confident that the model presents the best compromise as the basis of providing advice.

The B044 assessment indicates a long term decline in F since the late 80's with a corresponding increase in SSB since 1997. Recruitment is likely to have varied without much of a trend over most of this period although indicating a slight decline in recent years in conjunction with somewhat lower inter-annual variability.

Comparison with the continuity assessment was carried out on the basis of the WG proposed B043 model. The two assessments were found to be highly conservative in the estimation of stock dynamics and the panel agreed that the comparison with B043 was sufficient to assess the suitability of the B044 model. Given the similarity of the B043 and B044 (final model) it is unlikely that this perception will change. B044 scaled F estimates upwards and SSB estimates downwards slightly over the entire time-series. It suggested a less of a decline in recent recruitment than the continuity run.

During the assessment presentation and the discussions regarding the research priorities, a desire emerged to develop the current model into a fully-fledged SCA model by the incorporation of length information. I am not clear at to the motivation for this, i.e., what failure in the assessment is to be addressed. I mention this at this point, because to me it suggests that the WG believes that with more time they could have achieved a better assessment. I am not convinced that this would be the case for the following reasons:

Bluefish are rather fast growing, especially at young ages. This is one of the reasons that seasonal ALKs are used to age the catches. The length of a 4 year-old caught in Q1 is more like the length of a 3 year-old caught in Q4, in other words, without use of seasonal ALKs cohorts are likely to be smeared when using annual ALKs. Inevitably to maintain the same level of accuracy in a length-based assessment it is necessary to move to a semi-annual time step. Technically, this is not a problem, but with aggregated fleets and a seasonal pattern of migration one cannot assume that selectivity is constant within the seasonal cycle. It therefore becomes necessary to divide catches into gear and season combinations to establish consistent selectivities. At the low levels of aggregation, length information may not be sufficient to describe the average length distributions adequately. In short, going to a length based-assessment is not wrong, but it is not necessarily better, because to do it correctly will require in this case a significantly greater number of parameters to be estimated based on fewer data points.

Length based selectivities are to be preferred when highly selective gears such as gill nets interact with variable growth. For bluefish only a small portion of the catches is taken by highly selective gears and there is little evidence of high inter-annual variability in growth. The apparent age specificity in gears is much more likely to be associated with spatial availability of age groups rather than a gear effect by length.

I feel that the potential gains made in model realism will be outweighed by the increased uncertainty in the parameter estimation. Focusing on efforts to develop spatially specific models is likely to be of more value than pursuing a length based selection approach.

Given the uncertainty in a number of the factors considered within the assessment, to me it seems the potential uncertainty as assessed by the assessment is relatively small. The WG, based on a number of independent sensitivity analyses, suggested an uncertainty inflation factor of 15% was appropriate. Although numerically correct in terms of the management metrics in F and SSB, this estimate is based on individual sensitivities explored one at a time. The additive, or even interactive sensitivity, was not explored and is potentially greater.

5. **State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.**

The key elements of this ToR were addressed successfully by the WG: The WG examined the potential for estimating MSY-reference points and concluded, that although it was possible to numerically estimate such reference points, the contrast in SSB was insufficient to describe the stock-recruit relationship outside of the observed range of SSBs. The panel agreed with this conclusion and the subsequent decision to move to biomass reference points. Consistent with previous assessments the WG estimated a total stock biomass reference point, $TSB_{40\%}$, on the basis that exploitation is initiated prior to maturation. The intent of biomass reference points is not to describe the rate of exploitation, but to define the potential of the stock to reproduce. Therefore the panel considered a biomass reference point based on SSB to be more appropriate.

The appropriate choice of the biomass reference point (40% previously) was also considered by the panel. Forty percent appears to be a common value used in many North American demersal stocks. I am not aware of the exact basis for the value, but note that behaviour and life-history of bluefish are distinctly pelagic. Pelagic species are usually considered to be more productive and able to sustain higher rates of exploitation, so for bluefish a lower reference point may be appropriate. However, given the greater longevity and uncertainty in M, there was little basis for arguing for a more appropriate reference point. Consequently, $SSB_{40\%}$ was provided as the biomass reference point on the basis of consistency with previous recommendations rather than any specific conclusion regarding the appropriateness of 40%.

Using the accepted new ASAP stock assessment model B044, the SAW provided new recommended values for the reference points. The FMSYproxy = F40% = 0.170; BMSYproxy = SSB40% = 111,228 mt; $B_{THRESHOLD}$ = ½ SSB40% = 55,614 mt; and the proxy estimate for MSY = MSY40% = 13,967 mt. The WG did not provide estimates of uncertainty of the reference points.

6. **Evaluate stock status with respect to the existing model (from previous peer review accepted assessment) and with respect to a new model developed for this peer review.**

- **a) When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.**
- **b) Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).**

This ToR was addressed successfully by the WG:

a) The existing reference points are F_{MSY} = 0.19 and B_{MSY} = 147,052 MT (½ B_{MSY} = 73,526 MT) The 2014 F estimate (0.141) is well below F_{MSY} and the 2014 estimate of B is 92,755 MT, below B_{MSY} but well above ½ B_{MSY} overfishing is not occurring and that the stock is not overfished. Note that the biomass reference points here unusually refer to total biomass.

b) The state of the stock in relation to the new reference points is as follows:

$F_{MSYproxy} = 0.170 > 0.157 = F_{2014}$

$B_{MSYproxy} = 111{,}228$ mt $> 86{,}534$ mt $= SSB_{2014}$

$B_{THRESHOLD} = 55{,}614$ mt $< 86{,}534$ mt $= SSB_{2014}$

7. **Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level; see Appendix to the SAW TORs).**

a) **Provide annual projections (3 years). For given catches, each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).**
b) **Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.**
c) **Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.**

The key elements of this ToR were addressed successfully by the WG:

The SARC Panel agreed that the SAW met the key elements of this ToR.

a) Projections from the accepted model were made using standard NEFSC software. Projections for a range of constant F scenarios for 2016-2018 and 2015 removals set to the 2015 quota were run, all starting with initial abundance estimates drawn from MCMC runs of the accepted model B044. The projections were done using a single fleet; this required development of a combined selectivity curve, based on the last three years of the model estimates. A small CV was added to the selectivity-at-age estimates, as was also done for biological parameters drawn from lognormal distributions. The WG ran a number of sensitivity tests using model B043 to address potential uncertainties using the model as presented to the SARC; based on those tests, conclusions drawn using the accepted model B044 are expected to be robust. The ToR calls for the provision of annual probabilities of exceeding threshold reference points (below for SSB and above for F); these were not presented in respect to the final model (B044) but it is clear from the projected estimates of SSB and from the text provided by the WG that the probabilities in each of the years 2016-2018 of exceeding the biomass threshold is zero.
b) The WG considered the base model configuration the most realistic projection scenario. I did not agree with the WG that most recent recruitment estimates are more reliable than earlier ones as this statement does not consider the retrospective bias. Final projections did however use the whole time series as the value was very close to the average of the last 10 years.
c) I considers the projections appropriate for management given the WG's useful comments related to productivity and susceptibility of bluefish. The stock is unlikely to become overfished over the next three years if fishing is at or near the $F_{MSYproxy}$ and the projections support this view.

8. **Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports, as well as MAFMC SSC model recommendations from 2005 and the research recommendations contained in its 23 September 2013 report to the MAFMC. Identify new research recommendations.**

The progress against research recommendations was presented by the WG. The recommendations are set by various bodies, usually in response to previous assessments and their review. Frequently, they recommend a particular analysis or approach to be pursued, without specifically stating as to how this might be likely to improve the assessment in terms of its weaknesses. Some recommendations are based on reviews of other stocks, without a consideration of whether a similar approach may be useful for a specific species (for example the environmental correction of survey indices based on the available habitat niche). Lastly, some seem to relate to the expertise / beliefs of individual reviewers, who may well have a clear idea of how to use the recommended information, but it is not apparent how this fits in with the assessment process (for example conduct tagging).

The WG assisted the panel during the SARC meeting to provide more background information on a number of research recommendations. Based on this additional information the panel commented on the appropriateness in the consensus report.

With regards to future bluefish assessments, I am wondering if there is not the risk of too much effort being put into the move towards a length and age based SCA assessment. It seems to me that this is a generic type request that will result in substantial work with likely little gain in terms of improvement in advice for the reasons described in detail in under the ToR4 in my report.

Appendix 1: Review Committee members

Cynthia Jones
Normal Hall
Kevin Stokes
Sven Kupschus

Appendix 2: Terms of reference (from Annex 2, Statement of Work)

No changes to these were made prior to or during the SARC 48 review meeting.

A. Scup

1. Estimate catch from all sources including landings and discards. Include recreational discards, as appropriate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

2. Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). Characterize the uncertainty and any bias in these sources of data.

3. Describe the thermal habitat and its influence on the distribution and abundance of scup, and attempt to integrate the results into the stock assessment.

4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review.

a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

7. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) (see Appendix to SAW TORs for definitions).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

8. Review, evaluate and report on the status of the SARC, SSC, and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.

B. Bluefish

a. Estimate catch from all sources including landings and discards. Evaluate and if necessary update the discard mortality estimate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

b. Present and evaluate data and trends on life history information including, age, growth, natural mortality, food habits, and maturity.

c. Present the survey data available for use in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.), evaluate the utility of the age-length key for use in stock assessment, and explore standardization of fishery-independent indices. Investigate the utility of recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data, including exploring environmentally driven changes in availability and related changes in size structure. Explore the spatial distribution of the stock over time, and whether there are consistent distributional shifts.

d. Estimate relative fishing mortality, annual fishing mortality, recruitment, total abundance, and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Explore inclusion of multiple fleets in the model. Include both internal and historical retrospective analyses to allow a comparison with previous assessment results and previous projections. Explore alternative modeling approaches if feasible.

e. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

f. Evaluate stock status with respect to the existing model (from previous peer review accepted assessment) and with respect to a new model developed for this peer review.
    i. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
    ii. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

g. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level; see Appendix to the SAW TORs).
    i. Provide annual projections (3 years). For given catches, each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about

the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

ii. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

iii. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

h. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports, as well as MAFMC SSC model recommendations from 2005 and the research recommendations contained in its 23 September 2013 report to the MAFMC. Identify new research recommendations.

Appendix 3: Final Agenda

60th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for A. scup and B. bluefish

June 2-5, 2015

Stephen H. Clark Conference Room – Northeast Fisheries Science Center
Woods Hole, Massachusetts

Draft AGENDA* (version: April 30, 2015)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Tuesday, June 2

10 – 10:30 AM
Welcome — **James Weinberg**, SAW Chair
Introduction — **Cynthia Jones,** SARC Chair
Agenda
Conduct of Meeting

10:30 – 12:30 PM Assessment Presentation (A. Scup)
Mark Terceiro TBD **TBD**

12:30 – 1:30 PM Lunch

1:30 – 3:30 PM Assessment Presentation (A. Scup)
Mark Terceiro TBD **TBD**

3:30 – 3:45 PM Break

3:45 – 5:45 PM SARC Discussion w/ Presenters (A. Scup)
Cynthia Jones, SARC Chair **TBD**

5:45 – 6 PM Public Comments

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Wednesday, June 3

8:30 – 10:30 AM Assessment Presentation (B. Bluefish)

	Tony Wood	TBD	**TBD**
10:30 – 10:45 AM	Break		
10:45 – 12:30 PM	(cont.) Assessment Presentation (B. Bluefish)		
	Tony Wood	TBD	**TBD**
12:30 – 1:30 PM	Lunch		
1:30 – 3:30 PM	SARC Discussion w/presenters (B. Bluefish)		
	Cynthia Jones, SARC Chair		**TBD**
3:30 – 3:45 PM	Public Comments		
3:45 -4 PM	Break		
4 – 6 PM	Revisit with presenters (A. Scup)		
	Cynthia Jones, SARC Chair		**TBD**
7 PM	**(**Social Gathering **)**		

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Thursday, June 4

8:30 – 10:30	Revisit with presenters (B. Bluefish)		
	Cynthia Jones, SARC Chair		**TBD**
10:30 – 10:45	Break		
10:45 – 12:15	Review/edit Assessment Summary Report (A. Scup)		
	Cynthia Jones, SARC Chair		**TBD**
12:15 – 1:15 PM	Lunch		
1:15 – 2:45 PM	(cont.) edit Assessment Summary Report (A. Scup)		
	Cynthia Jones, SARC Chair		**TBD**
2:45 – 3 PM	Break		
3 – 6 PM	Review/edit Assessment Summary Report (B. Bluefish)		
	Cynthia Jones, SARC Chair		**TBD**

Friday, June 5

9:00 AM – 5:00 PM SARC Report writing.

*All times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public. During the SARC report writing stage on June 5, the public should not engage in discussion with the SARC.

Appendix 4: Bibliography provided

Scup

Conn, Paul B. 2010. Hierarchical Analysis of Multiple Noisy Abundance Indices. Canadian Journal of Fisheries and Aquatic Sciences: 67. 108-120. 13p.

Love, J.W. and Chase, P.D. 2009. Geometric Morphological Differences Distinguish Populations of Scup in the Northwestern Atlantic Ocean. Marine and Coastal Fisheries: Dynamics, Management and Ecosystem Science 1:1. 22-28. 7p.

Mayo, R.K. 1983. An Assessment of the Scup *Stenotomus chrysops* Population in the Southern New England and Middle Atlantic Regions. Woods Hole Laboratory Reference Doc 82-46. 68p.

Miller, T. et al. January 2009. Report by the Peer Review Panel for the Northeast Data Poor Stocks Working Group. 38p.

Northeast Fisheries Science Center (NEFSC). 2008. Northeast Data Poor Stocks Working Group Report. NEFSC Reference Document 09-02A & B. 886p.

Northeast Fisheries Science Center (NEFSC). 2012. Stock Assessment of Scup (*Stenotomus chrysops*) for 2012. NEFSC Reference Document 12-25. 112p.

NOAA Fisheries Toolbox (NFT). April 2014. A Stroll Through the Automated Plots for ASAP3. Technical Documentation for ASAP Toolbox Version 3. 6p.

NOAA Fisheries Toolbox (NFT). September 2012. Technical Documentation for ASAP Version 3.0. 71p.

NOAA Fisheries Toolbox (NFT). September 2012. User Manual for ASAP 3. 22p.

Working Papers

Scup Working Group. SAW-SARC 60 Stock Assessment of Scup for 2015. 303p.

Bell, R. J., Gervelis, B. and Hoey, J. Northeast Fisheries Science Center (NEFSC), National Marine Fisheries Service. April 2015. Study Fleet Scup CPUE. SAW-SARC Background Paper A12. 10p.

Lapp, M., Almeida, K., Cadrin, S. 2015. Commercial Fishermen's Perspective on the Scup Fishery and Resource. SAW-SARC 60 Background Paper A18. 6p.

Manderson, J. et al. Northeast Fisheries Science Center (NEFSC) Cooperative Research Program and Virginia Institute of Marine Science (VIMS). 2015. Description of the Thermal Habitat and its Influence on the Distribution and Abundance of Scup, and Attempt to Integrate the Results Into the Stock Assessment (TOR 3). SAW-SARC 60 Background Paper A11. 52p.

Terceiro, M. 2015. Biological Data for Scup (TOR 2 and 3). SAW-SARC 60 Background Paper A2. Northeast Fisheries Science Center (NEFSC). 47p.

Terceiro, M. 2015. Consideration of Alternative Strata Sets and Model-Based Indices from NEFSC Spring and Fall Trawl Surveys for Scup (TOR 2). SAW-SARC 60 Background Paper A4. 48p.

Terceiro, M. 2015. Description of Commercial Fishery Dealer Report Trawl Gear Landings and Effort and Modeling Landings Rate (LPUE) Data for Scup (TOR 1). SAW-SARC 60 Background Paper A5. 34p.

Terceiro, M. 2015. Description of MRFSS/MRIP Intercept Catch and Effort and Modeling of Total Catch Rate (CPUE) Data for Scup (TOR 1). SAW-SARC 60 Background Paper A9. 13p.

Terceiro, M. 2015. An Approach to the Integration of Survey Abundance Indices Used for Scup Population Model Calibration (TOR 2 and 8). SAW-SARC 60 Background Paper A10. 48p.

Terceiro, M. and Miller, A. 2015. Estimates of Commercial Fishery Scup Discards: 1989-2013 (TOR1). SAW-SARC 60 Background Paper A3. 37p.

Terceiro, M. and Miller, A. 2015. Description of Vessel Trip Report Trawl Gear Catch and Effort Data and Modeling Catch Rates (CPUE) for Scup (TOR 1). SAW-SARC 60 Background Paper A6. 27p.

Terceiro, M. and Miller, A. 2015. Description of Vessel Trip Report Party/Charter Boat Catch and Effort Data and Modeling Catch Rates (CPUE) for Scup (TOR 1). SAW-SARC 60 Background Paper A7. 29p.

Terceiro, M. and Miller, A. 2015. Description and Modeling of NEFOP (Observer) Fish Trawl Gear Catch Rate (CPUE) Data for Scup (TOR 1). SAW-SARC 60 Background Paper A8. 23p.

Presentations

Conn, P. Integrated Indices for Scup. SAW-SARC Background Paper A13. PowerPoint presentation. 15 slides.

NEFSC Trawl Surveys: Spatial and Temporal Distribution of Catch, Bottom Temperature and Salinity. SAW-SARC Background Paper A14. PowerPoint presentation. 83 slides.

NEFSC Trawl Surveys: Spatial and Temporal Distribution of Catch and Bottom Temperature V2: Revised Catch Intervals, 'Bubble' Size, and to Include Strata Boundaries. SAW-SARC Background Paper A15. PowerPoint presentation. 48 slides.

Spatial and Temporal Distribution of Fishery Catch and Effort. SAW-SARC 60 Background Paper A16. PowerPoint presentation. 22 slides.

Bluefish

ASMFC Technical Committee/Assessment Subcommittee. 2005. 41st SAW Assessment Report of Bluefish. 85p.

Conn, Paul B. 2010. Hierarchical Analysis of Multiple Noisy Abundance Indices. Canadian Journal of Fisheries and Aquatic Sciences: 67: 108-120. 13p.

Dick, E.J. and MacCall, A.D. 2011. Depletion-Based Stock Reduction Analysis: A Catch-Based Method for Determining Sustainable Yields for Data-Poor Fish Stocks. Journal of Fisheries Research: 110: 331-341. 11p.

Jones, C.M. 2005. 41st Northeast Regional Stock Assessment Workshop (SAW-41) Chair's Report for Peer Review. 89p.

MacCall, A.D. 2009. Depletion-Corrected Average Catch: A Simple Formula for Estimating Sustainable Yields in Data-Poor Situations. ICES Journal of Marine Sciences, 66: 2267-2271. 5p.

NOAA. 2006. Bluefish, *Pomatomus saltatrix*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-198. 100 p.

Robillard, E., Reiss, C.S. and Jones, C.M. 2008. Reproductive Biology of Bluefish (*Pomatomus saltatrix*) Along the East Coast of the United States. Journal of Fisheries Research: 90: 198-208. 11p.

Robillard, E., Reiss, C.S. and Jones, C.M. 2009. Age-Validation and Growth of Bluefish (*Pomatomus saltatrix*) Along the East Coast of the United States. Journal of Fisheries Research: 95: 65-75. 11p.

Salerno, D.J., Burnett, J. and Ibara, R.M. 2001. Age, Growth, Maturity and Spatial Distribution of Bluefish, *Pomatomus saltatrix* (Linnaeus), off the Northeast Coast of the United States, 1985-96. Journal of Northwest Atlantic Fisheries Science: 29: 31-39. 9p.
Shepherd, G.R. et al. 2006. The Migration Patterns of Bluefish (*Pomatomus saltatrix*) Along the Atlantic Coast, Determined from Tag Recoveries. Fishery Bulletin 104: 559-570. 12p.

Working Papers

Bluefish Working Group. SAW-SARC 60 Stock Assessment of Bluefish for 2015. 399p.
Bluefish Working Group. 2015. Life History of Bluefish (TOR 2). SAW-SARC 60 Background Paper B2. 32p.
Bluefish Working Group. 2015. Improvements to Ageing Data for the Atlantic Coast Bluefish Stock. SAW-SARC 60 Background Paper B5. 92p.
Bluefish Working Group. 2015. Age Data Decisions Made at Assessment Workshop. SAW-SARC 60 Background Paper B6. 15p.
Celestino, M. and Brust, J. 2015. Implications of Expanded Age Sampling on Bluefish *Pomatomus saltatrix* Catch Estimates. SAW-SARC 60 Background Paper B7. 33p.
Gottschall, K. 2015. Review and Evaluation of Catch-and-Release Angling Mortality Estimates for Bluefish. SAW-SARC 60 Background Paper B1. 15p.
Manderson, J. et al. 2015. Exploring Temperature Effects on the Distribution of Bluefish and Availability to Surveys. SAW-SARC 60 Background Paper B4. 45p.
Spanik, K., Gartland, J., Gaichas, S. 2015. Food Habitats of Bluefish (TOR 2). SAW-SARC 60 Background Paper B3. 21p.

Appendix 5:Statement of Work

Statement of Work

60th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for scup and bluefish

Statement of Work (SOW) for CIE Panelists
(including a description of SARC Chairman's duties)

BACKGROUND

The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Representative (COR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are independently selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

SCOPE

Project Description: The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC peer review is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development and report preparation (which is done by SAW Working Groups or ASMFC technical committees), assessment peer review (by the SARC), public presentations, and document publication. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fisheries within the jurisdiction of NOAA's Greater Atlantic Regional Fisheries Office (GARFO).

The purpose of this meeting will be to provide an external peer review of benchmark stock assessments for **scup** and **bluefish**.

OBJECTIVES

The SARC review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the SSC of the New England or

Mid-Atlantic Fishery Management Council. The SARC panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report.

Duties of reviewers are explained below in the "**Requirements for CIE Reviewers**", in the "**Charge to the SARC Panel**" and in the "**Statement of Tasks**". The draft stock assessment Terms of Reference (ToRs) which are carried out by the SAW WGs are attached in **Annex 2**. The draft agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is described in **Annex 4.**

Requirements for the reviewers: Three reviewers shall conduct an impartial and independent peer review of the **scup** and **bluefish** stock assessments, and this review should be in accordance with this SoW and stock assessment ToRs herein. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include statistical catch-at-age, state-space and index models. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points that includes an appreciation for the varying quality and quantity of data available to support estimation of Biological Reference Points. SARC 59 will address fishery stock assessments of **scup** and **bluefish**. For both species, experience in assessing pelagic stocks and in incorporating environmental factors into assessments would be desirable. For bluefish, experience in the use of recreational fisheries data would also be desirable.

PERIOD OF PERFORMANCE

The contractor shall complete the tasks and deliverables as specified in the schedule of milestones within this statement of work. Each reviewer's duties shall not exceed a maximum of 16 days to complete all work tasks of the peer review described herein.

Not covered by the CIE, the SARC chair's duties should not exceed a maximum of 16 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the open meeting for SARC Summary Report preparation).

PLACE OF PERFORMANCE AND TRAVEL

Each reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during June 2-5, 2015.

STATEMENT OF TASKS

Charge to SARC panel: During the SARC meeting, the panel is to determine and write down whether each stock assessment Term of Reference (ToR) of the SAW (see **Annex 2**) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. **If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.** Where possible, the SARC chair shall identify or

facilitate agreement among the reviewers for each stock assessment Term of Reference of the SAW.

If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and $MSY_{)}$, the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.

Each reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Tasks prior to the meeting: The contractor shall independently select qualified reviewers that do not have conflicts of interest to conduct an independent scientific peer review of stock assessments prepared by SAW WGs or ASMFC Technical Committees in accordance with the tasks and ToRs within the SoW. Upon completion of the independent reviewer selection by the contractor's technical team, the contractor shall provide the reviewer information (full name, title, affiliation, country, address, email, FAX number, and CV suitable for public distribution) to the COR, who will forward this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The contractor shall be responsible for providing the SoW and stock assessment ToRs to each reviewer. The NMFS Project Contact will be responsible for providing the reviewers with the background documents, reports for review, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact will also be responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COR prior to the commencement of the peer review.

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX or by email the following requested information (e.g., 1.name [first, middle, and last], 2.contact information, 3.gender, 4.country of birth, 5.country of citizenship, 6.country of permanent residence, 7.whether there is dual citizenship, 8.country of current residence, 9.birth date [mo, day, year], 10.passport number, 11.country of passport) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/.

Pre-review Background Documents and Working Papers: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the SARC chair and CIE reviewers the necessary background information and reports (i.e., working papers) for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

Tasks during the panel review meeting: Each reviewer shall conduct the independent peer review of the stock assessments in accordance with the SoW and stock assessment ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and contractor.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the stock assessment ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

(SARC chair)
Act as chairperson, where duties include control of the meeting, coordination of presentations and discussions, making sure all stock assessment Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For each assessment, review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to discuss the stock assessment and to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(SARC CIE reviewers)
For each stock assessment, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a reviewer's point of view, determine whether each stock assessment Term of Reference of the SAW was completed successfully. Terms of Reference that are completed successfully are likely to serve as a basis for providing scientific advice to management. If a reviewer considers any existing Biological Reference Point or BRP proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist. Review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

Tasks after the panel review meeting:

SARC CIE reviewers:

Each CIE reviewer shall prepare an Independent CIE Report (see **Annex 1**). This report should explain whether each stock assessment Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified above in the "Charge to SARC panel" statement.

If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent CIE Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.

During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent CIE Report produced by each reviewer.

The Independent CIE Report can also be used to provide greater detail than the SARC Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

SARC chair:
The SARC chair shall prepare a document summarizing the background of the work to be conducted as part of the SARC process and summarizing whether the process was adequate to complete the stock assessment Terms of Reference of the SAW. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the SARC Summary Report (see **Annex 4**).

SARC chair and CIE reviewers:
The SARC Chair, with the assistance from the CIE reviewers, will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each stock assessment Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this SARC Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion.

The SARC Summary Report (please see **Annex 4** for information on contents) should address whether each stock assessment Term of Reference of the SAW was completed successfully. For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. The Report should also include recommendations that might improve future assessments.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

The contents of the draft SARC Summary Report will be approved by the CIE reviewers by the end of the SARC Summary Report development process. The SARC chair will complete all final editorial and formatting changes prior to approval of the contents of the draft SARC Summary Report by the CIE reviewers. The SARC chair will then submit the approved SARC Summary Report to the NEFSC contact (i.e., SAW Chairman).

DELIVERY

Each reviewer shall complete an independent peer review report in accordance with the SoW. Each reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each reviewer shall complete the independent peer review addressing each stock assessment ToR listed in **Annex 2**.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting at the Woods Hole, Massachusetts scheduled during the tentative dates of June 2-5, 2015.
3) Conduct an independent peer review in accordance with this SoW and the assessment ToRs (listed in **Annex 2**).
4) No later than June 19, 2015, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator, via email to mshivlani@ntvifederal.com, and to Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each assessment ToR in **Annex 2**.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

April 24, 2015	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
May 19, 2015	NMFS Project Contact will attempt to provide reviewers the pre-review documents
June 2-5, 2015	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA

June 5, 2015	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
June 19, 2015	Reviewers submit draft independent peer review reports to the contractor's technical team for independent review
June 19, 2015	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
June 26, 2015	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
July 2, 2015	Contractor submits independent peer review reports to the COR who reviews for compliance with the contract requirements
July 10, 2015	The COR distributes the final reports to the NMFS Project Contact and regional Center Director

* The SARC Summary Report will not be submitted, reviewed, or approved by the CIE.

The SAW Chairman will assist the SARC chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion.

NEFSC staff and the SAW Chairman will make the final SARC Summary Report available to the public. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers, which will serve as a SAW Assessment Report.

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on substitutions. The COR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: The deliverables shall be the final peer review report from each reviewer that satisfies the requirements and terms of reference of this SoW. The contract shall be successfully completed upon the acceptance of the contract deliverables by the COR based on three performance standards:

(1) each report shall be completed with the format and content in accordance with **Annex 1**,
(2) each report shall address each stock assessment ToR listed in **Annex 2**,
(3) each report shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Upon the acceptance of each independent peer review report by the COR, the reports will be distributed to the NMFS Project Contact and pertinent NMFS science director, at which time the reports will be made publicly available through the government's website.

The contractor shall send the final reports in PDF format to the COR, designated to be Allen Shimada, via email allen.shimada@noaa.gov

Support Personnel:

Allen Shimada, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
allen.shimada@noaa.gov Phone: 301-427-8174

Manoj Shivlani, CIE Lead Coordinator
NTVI Communications, Inc.
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-968-7136

Key Personnel:

Dr. James Weinberg, NEFSC SAW Chairman, NMFS Project Contact
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov (Phone: 508-495-2352) (FAX: 508-495-2230)

Dr. William Karp, NEFSC Science Director
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
william.karp@noaa.gov Phone: 508-495-2233

Annex 1: Format and Contents of Independent Peer Review Report

1. The independent peer review report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The main body of the report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Findings of whether they accept or reject the work that they reviewed, and an explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each ToR, and Conclusions and Recommendations in accordance with the ToRs. For each assessment reviewed, the report should address whether each ToR of the SAW was completed successfully. For each ToR, the Independent Review Report should state why that ToR was or was not completed successfully. To make this determination, the SARC chair and reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the SARC Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The independent report shall be a stand-alone document for others to understand the proceedings and findings of the meeting, regardless of whether or not others read the SARC Summary Report. The independent report shall be an independent peer review of each ToR, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of this Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: 60th SAW/SARC Stock Assessment Terms of Reference (file vers.: 10/162014)

C. Scup

1. Estimate catch from all sources including landings and discards. Include recreational discards, as appropriate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

2. Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). Characterize the uncertainty and any bias in these sources of data.

3. Describe the thermal habitat and its influence on the distribution and abundance of scup, and attempt to integrate the results into the stock assessment.

4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review.

a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

7. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) (see Appendix to SAW TORs for definitions).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

8. Review, evaluate and report on the status of the SARC, SSC, and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.

Annex 2: (cont)

D. Bluefish

a. Estimate catch from all sources including landings and discards. Evaluate and if necessary update the discard mortality estimate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

b. Present and evaluate data and trends on life history information including, age, growth, natural mortality, food habits, and maturity.

c. Present the survey data available for use in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.), evaluate the utility of the age-length key for use in stock assessment, and explore standardization of fishery-independent indices. Investigate the utility of recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data, including exploring environmentally driven changes in availability and related changes in size structure. Explore the spatial distribution of the stock over time, and whether there are consistent distributional shifts.

d. Estimate relative fishing mortality, annual fishing mortality, recruitment, total abundance, and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Explore inclusion of multiple fleets in the model. Include both internal and historical retrospective analyses to allow a comparison with previous assessment results and previous projections. Explore alternative modeling approaches if feasible.

e. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

f. Evaluate stock status with respect to the existing model (from previous peer review accepted assessment) and with respect to a new model developed for this peer review.
    i. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
    ii. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

g. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level; see Appendix to the SAW TORs).
    i. Provide annual projections (3 years). For given catches, each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about

the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

ii. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

iii. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

h. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports, as well as MAFMC SSC model recommendations from 2005 and the research recommendations contained in its 23 September 2013 report to the MAFMC. Identify new research recommendations.

Annex 2: (cont)

Appendix to the SAW Assessment TORs:
Clarification of Terms used in the SAW/SARC Terms of Reference

On "Overfishing Limit" and Acceptable Biological Catch" (DOC Nat. Stand. Guidel. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of [overfishing limit] OFL and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect annual catch that is consistent with schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, [optimal yield] OY does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce MSY and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Interactions among members of a SAW Assessment Working Group:

Anyone participating in SAW assessment working group meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

One model or alternative models:

The preferred outcome of the SAW/SARC is to identify a single "best" model and an accompanying set of assessment results and a stock status determination. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility

each model, including a comparison of results.

Annex 3: Draft Agenda

60th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for A. scup and **B. bluefish**

June 2-5, 2015

Stephen H. Clark Conference Room – Northeast Fisheries Science Center
Woods Hole, Massachusetts

DRAFT AGENDA* (version: Dec. 1, 2014)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Tuesday, June 2

10 – 10:30 AM

Welcome	**James Weinberg**, SAW Chair	
Introduction	**TBD,** SARC Chair	
Agenda		
Conduct of Meeting		

10:30 – 12:30 PM	Assessment Presentation (A. Scup)		
	Mark Terceiro	TBD	**TBD**
12:30 – 1:30 PM	Lunch		
1:30 – 3:30 PM	Assesssment Presentation (A. Scup)		
	Mark Terceiro	TBD	**TBD**
3:30 – 3:45 PM	Break		
3:45 – 5:45 PM	SARC Discussion w/ Presenters (A. Scup)		
	TBD, SARC Chair		**TBD**
5:45 – 6 PM	Public Comments		

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Wednesday, June 3

8:30 – 10:30 AM	Assessment Presentation (B. bluefish)		
	Tony Wood	TBD	**TBD**
10:30 – 10:45 AM	Break		
10:45 – 12:30 PM	(cont.) Assessment Presentation (B. bluefish)		
	Tony Wood	TBD	**TBD**
12:30 – 1:30 PM	Lunch		
1:30 – 3:30 PM	SARC Discussion w/presenters (B. bluefish)		
	TBD, SARC Chair		**TBD**
3:30 – 3:45 PM	Public Comments		
3:45 -4 PM	Break		
4 – 6 PM	Revisit with presenters (A. Scup)		
	TBD, SARC Chair		**TBD**
7 PM	**(**Social Gathering **)**		

TOPIC PRESENTER(S) SARC LEADER RAPPORTEUR

Thursday, June 4

8:30 – 10:30	Revisit with presenters (B. bluefish) **TBD,** SARC Chair	**TBD**
10:30 – 10:45	Break	
10:45 – 12:15	Review/edit Assessment Summary Report (A. Scup) **TBD,** SARC Chair	**TBD**
12:15 – 1:15 PM	Lunch	
1:15 – 2:45 PM	(cont.) edit Assessment Summary Report (A. Scup) **TBD,** SARC Chair	**TBD**
2:45 – 3 PM	Break	
3 – 6 PM	Review/edit Assessment Summary Report (B. bluefish) **TBD,** SARC Chair	**TBD**

Friday, June 5

9:00 AM – 5:00 PM	SARC Report writing.

*All times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public. The public should not engage in discussion with the SARC during SARC report writing, which is scheduled for June 5.

***The NMFS Project contact will provide the final agenda about four weeks before meeting.**
Reviewers must attend the entire meeting.

Annex 4: Contents of SARC Summary Report

1.

The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background, a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment reviewed, the report should address whether each Term of Reference of the SAW Working Group was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

To make this determination, the SARC chair and CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If the CIE reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

The report may include recommendations on how to improve future assessments.

2.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3.

The report shall also include the bibliography of all materials provided during the SAW, and relevant papers cited in the SARC Summary Report, along with a copy of the CIE Statement of Work.

The report shall also include as a separate appendix the assessment Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.